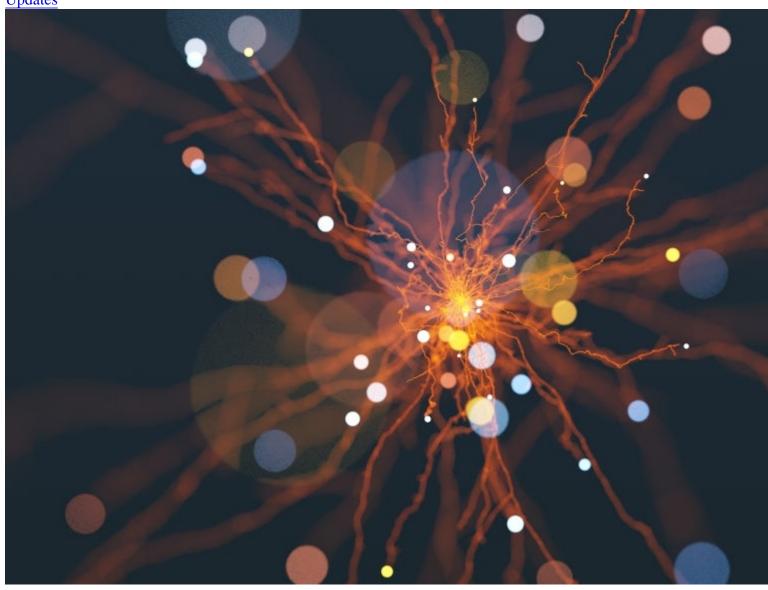In the last 30 days, California Governor Gavin Newsom signed 17 artificial intelligence (AI) bills and vetoed AI safety bill SB 1047. One of the AI-related bills signed into law, AB 2013 "Generative Artificial Intelligence: Training Data Transparency," imposes new disclosure requirements on the developers of generative artificial intelligence (GenAI) systems and services that are made available to Californians.

According to AB 2013's primary sponsor, Assemblymember Jacqui Irwin, AB 2013 is intended to allow consumers to "better evaluate if they have confidence in the AI system or service, compare competing systems and services, or put into place mitigation measures to address any shortcomings of the particular system or service." To this end, the bill imposes specific transparency requirements regarding the provenance and nature of training data.

Key Provisions of AB 2013

AB 2013 mandates that the developers of GenAI systems or services post detailed documentation on their websites about the data used to train these systems.

Applicability

The law applies to "developers," which is broadly defined as those who design, code, produce, or substantially modify an artificial intelligence system or service for use by members of the public. "Substantial modification" includes new versions, releases, or other updates that materially change the system or service's functionality or performance, including updates incorporating the results of retraining or fine-tuning of the model. The bill applies regardless of whether the system or service is provided for a fee.

The documentation requirements apply to developers of GenAI systems and services released or substantially modified on or after January 1, 2022.

Documentation Requirements

Prior to a GenAI system or service being released or substantially modified, the developer must post a "high-level" summary of the data used to train the system, including, at minimum:

1. The sources or owners of the datasets.
2. A description of how these datasets further the intended purpose of the system or service.
3. The number of data points included in the datasets, as a general range or with an estimated figure for dynamic datasets.
4. A description of the types of "data points" within the datasets, meaning either the types of labels used or, in datasets without labeling, the general characteristics of the datasets.
5. Whether the datasets included any data protected by copyright, trademark, or patent or are entirely in the public domain, and whether the datasets were purchased or licensed by the developer.
6. Whether the datasets included personal information or aggregate consumer information.
7. Whether there was any cleaning, processing, or other modification to the datasets by the developer and the intended purpose of those efforts.
8. The time period during which the data in the datasets were collected and a notice if the data collection is ongoing.
9. The dates the datasets were first used during the development of the system or service.
10. Whether the system or service used or continuously uses "synthetic data generation" in its development, including a description of the functional need or desired purpose of the synthetic data in relation to the intended purpose of the system or service.

Exemptions

The law exempts GenAI systems and services that (1) are used solely to ensure system security and integrity, (2) are used solely to operate aircraft in national airspace, and (3) are developed for national security, military, or defense purposes and are made available only to a federal entity.

Enforcement

AB 2013 makes no mention of enforcement or financial penalties for noncompliance. However, in an assembly committee on privacy and consumer protection hearing on the bill, the [supplied analysis](#) stated that the bill would likely be enforced under California's Unfair Competition law.

The bill's provisions come into effect on January 1, 2026.

Moving Forward

AB 2013 may be challenged in court on constitutional or other grounds, but if it comes into effect, the passage of AB 2013 has several implications for covered developers:

Retrospective Documentation

Because many GenAI models have training sets developed over time or have been trained on varying training sets over time, the law potentially requires retrospective documentation be compiled in order for a developer to be able to post on January 1, 2026, the required information about "*the data* used by the developer to train the generative artificial intelligence system or service" (emphasis added). Compiling retrospective information may take substantial time and effort, especially if the training data was not retained or well documented.

No Exemptions for Trade Secrets or Intellectual Property

AB 2013 contains no exemptions or limitations in the required information for trade secrets or intellectual property. This omission is unusual and, if this aspect of the law goes into effect, will likely present a difficult compliance challenge for covered GenAI developers.

Impact on Companies Specializing in Retraining or Integrating

Because the law covers those who substantially modify existing systems, businesses that specialize in retraining existing GenAI models or who integrate their own products with existing systems may be affected by the bill's requirements.

Authors

[Marc S. Martin](#)

Partner
MMartin@perkinscoie.com   [202.654.6351](#)

Sydney Veatch

Associate
SVeatch@perkinscoie.com

Explore more in

Technology Transactions & Privacy Law   Artificial Intelligence & Machine Learning

Related insights

Update

Implications of California Governor Newsom's Veto of AI Safety Bill SB 1047